Postulates for a Poisson Process

1. Events in non-overlapping time intervals are <u>independent</u>.
2. For "small" Δ_t :
   a) Pr [1 arrival in Δ_t] $\approx \lambda\Delta_t$
   b) Pr [more than 1 arrival in Δ_t] ≈ 0
   c) Pr [no arrivals in Δ_t] $\approx 1-\lambda\Delta_t$

Poisson Process

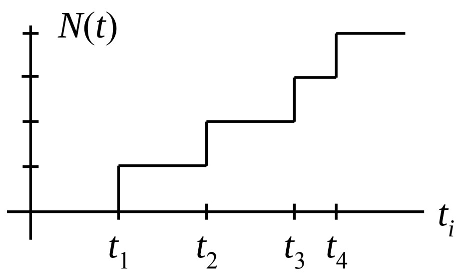
Arrivals:

Let $N(t_1,t_2)$ be the number of arrivals in the interval $[t_1, t_2)$

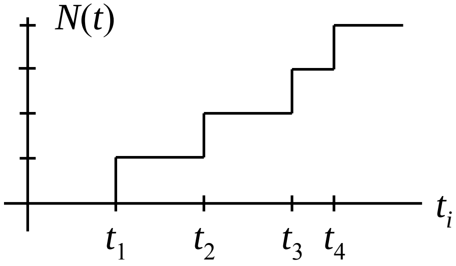
Let $N(t)$ be the number of arrivals in the interval $[0, t)$, i.e., $N(t) = N(0,t)$

$$\Pr\left[N(t)=k\right]=\frac{(\lambda t)^k}{k!}e^{-\lambda t}, \qquad k=0,1,2,\cdots$$

- Homogeneous in time: $N(t) = N(0, t) = \mathrm{N}(t_1, t_2)$, where $t = t_2 - t_1$



Poisson Process (cont'd.)



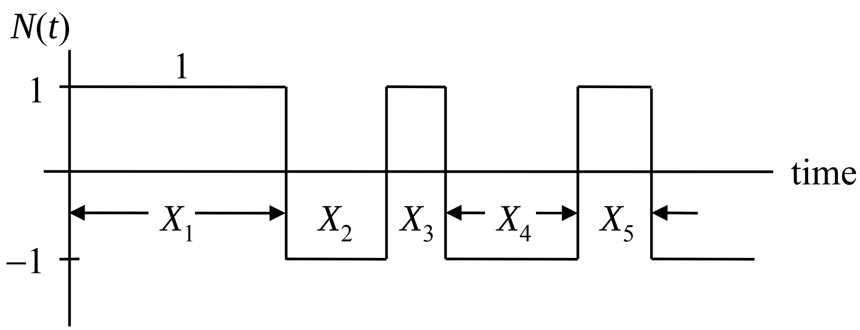
- If $[t_1, t_2]$ and $[t_3, t_4]$ are non-overlapping, $N(t_1, t_2)$ and $N(t_3, t_4)$ are independent.
- For $t_2 > t_1$, $N(t_1, t_2) = N(0, t_2) - N(0, t_1)$
- Mean, autocovariance, and autocorrelation functions follow (see text)

$$m_N(t) = E\left[N(t)\right] = \lambda t$$

$$C_N(t_1, t_0) = E\left[\left(N(t_1) - m_N(t_1)\right)\left(N(t_0) - m_N(t_0)\right)\right] = \lambda \min(t_1, t_0)$$

$$R_N(t_1, t_0) = C_N(t_1, t_0) + m_N(t_1) m_N(t_0) = \lambda \min(t_1, t_0) + \lambda^2 t_1 t_0$$

Random Telegraph Signal



$N(t)$ changes sign with each arrival of the Poisson process of rate λ

$$\Pr\left[N(0)=1\right]=p, \quad \Pr\left[N(0)=-1\right]=1-p$$

PMF of $N(t)$

$$\Pr\left[N(t)=1\right]=\Pr\left[N(t)=1\middle|N(0)=1\right]\Pr\left[N(0)=1\right]$$
$$+\Pr\left[N(t)=1\middle|N(0)=-1\right]\Pr\left[N(0)=-1\right]$$

$$\Pr\left[N(t)=-1\right]=\Pr\left[N(t)=-1\middle|N(0)=1\right]\Pr\left[N(0)=1\right]$$
$$+\Pr\left[N(t)=-1\middle|N(0)=-1\right]\Pr\left[N(0)=-1\right]$$

Even number by events:

$$\Pr\left[N(t)=\pm1\middle|N(0)=\pm1\right]=\sum_{k=0}^{\infty}\frac{(\lambda t)^{2k}}{(2k)!}e^{-\lambda t}=e^{-\lambda t}\left(\frac{e^{\lambda t}+e^{-\lambda t}}{2}\right)=\frac{1}{2}\left(1+e^{-2\lambda t}\right)$$

Odd number of events:

$$\Pr\left[N(t)=\pm1\middle|N(0)=\mp1\right]=\sum_{k=0}^{\infty}\frac{(\lambda t)^{2k+1}}{(2k+1)!}e^{-\lambda t}=e^{-\lambda t}\left(\frac{e^{\lambda t}-e^{-\lambda t}}{2}\right)=\frac{1}{2}\left(1-e^{-2\lambda t}\right)$$

PMF of $N(t)$ (continued)

We then have

$$\Pr\left[N(t)=1\right]=\frac{1}{2}\left(1+e^{-2\lambda t}\right)p+\frac{1}{2}\left(1-e^{-2\lambda t}\right)(1-p)$$

$$=\frac{p}{2}+\frac{p}{2}e^{-2\lambda t}+\frac{1}{2}-\frac{p}{2}-\frac{1}{2}e^{-2\lambda t}+\frac{p}{2}e^{-2\lambda t}=\frac{1}{2}+\left(p-\frac{1}{2}\right)e^{-2\lambda t}$$

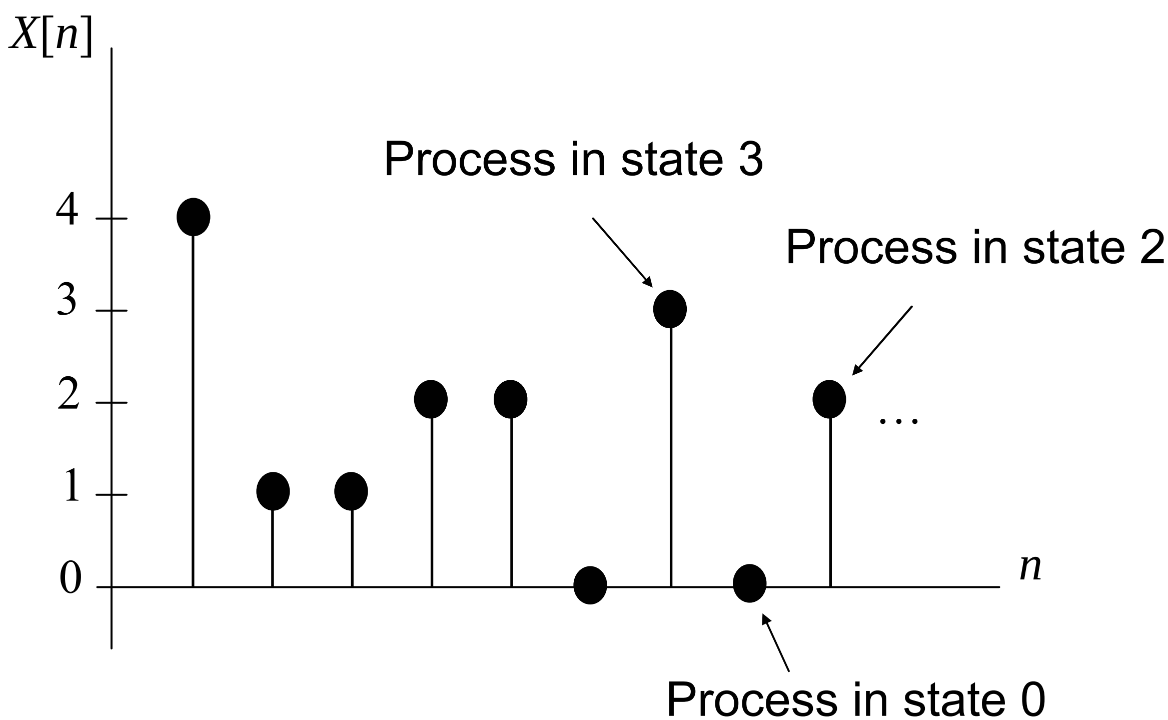
$$\Pr\left[N(t)=-1\right]=\frac{p}{2}\left(1-e^{-2\lambda t}\right)+\frac{1-p}{2}\left(1+e^{-2\lambda t}\right)=\frac{1}{2}-\left(p-\frac{1}{2}\right)e^{-2\lambda t}$$

Mean, autocorrelation, and autocovariance functions

$$m_N(t) = E\left[N(t)\right] = (-1)\Pr\left[N(t) = -1\right] + (1)\Pr\left[N(t) = 1\right]$$
$$= -\frac{1}{2} + \left(p - \frac{1}{2}\right)e^{-2\lambda t} + \frac{1}{2} + \left(p - \frac{1}{2}\right)e^{-2\lambda t} = (2p-1)e^{-2\lambda t}$$

$$R_N(t_1,t_2) = E\left[N(t_1)N(t_2)\right]$$
$$= (+1)\Pr\left[N(t_1) = N(t_2)\right] + (-1)\Pr\left[N(t_1) \neq N(t_2)\right]$$
$$= \frac{1}{2}\left[1 + e^{-2\lambda|t_2-t_1|}\right] - \frac{1}{2}\left[1 - e^{-2\lambda|t_2-t_1|}\right] = e^{-2\lambda|t_2-t_1|}$$
$$C_N(t_1,t_2) = R_N(t_1,t_2) = m_N(t_1)m_N(t_2) = e^{-2\lambda|t_2-t_1|} - (2p-1)^2 e^{-2\lambda(t_1+t_2)}$$

Markov Processes: Discrete-Time Markov Chain



Discrete-Time Markov Chain

Let X[n] be a discrete-time discrete-magnitude random signal. If it satisfies

$$\Pr\left[X[n]=j \,\middle|\, X[n-1]=i_1, X[n-2]=i_2,\cdots,X[1]=i_{n-1},X[0]=i_n\right]$$
$$=\Pr\left[X[n]=j \,\middle|\, X[n-1]=i_1\right] \qquad \text{for all } n, j, i_1, i_2,\cdots,i_n$$

then X[n] is called a discrete-time Markov chain.

State transition probabilities:

$$p_{ij}=\Pr\left[X[n]=j \,\middle|\, X[n-1]=i\right],$$
$$0\le i, j\le m-1$$
$$0\le p_{ij}\le 1; \quad \sum_{j=0}^{m-1} p_{ij}=1,$$
$$\text{for} \quad i=0,1,\cdots,m-1$$

State transition matrix:

$$\mathbf{P}=\begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots & p_{0,m-1} \\ p_{10} & p_{11} & p_{12} & \cdots & p_{1,m-1} \\ \vdots & \vdots & \vdots & & \vdots \\ p_{m-1,0} & p_{m-1,1} & p_{m-1,2} & \cdots & p_{m-1,m-1} \end{bmatrix}$$

Rows sum to 1

Example: Consider a 3-state Markov chain

$$\{-1,\ 0,\ +1\}$$

state 0 state 1 state 2

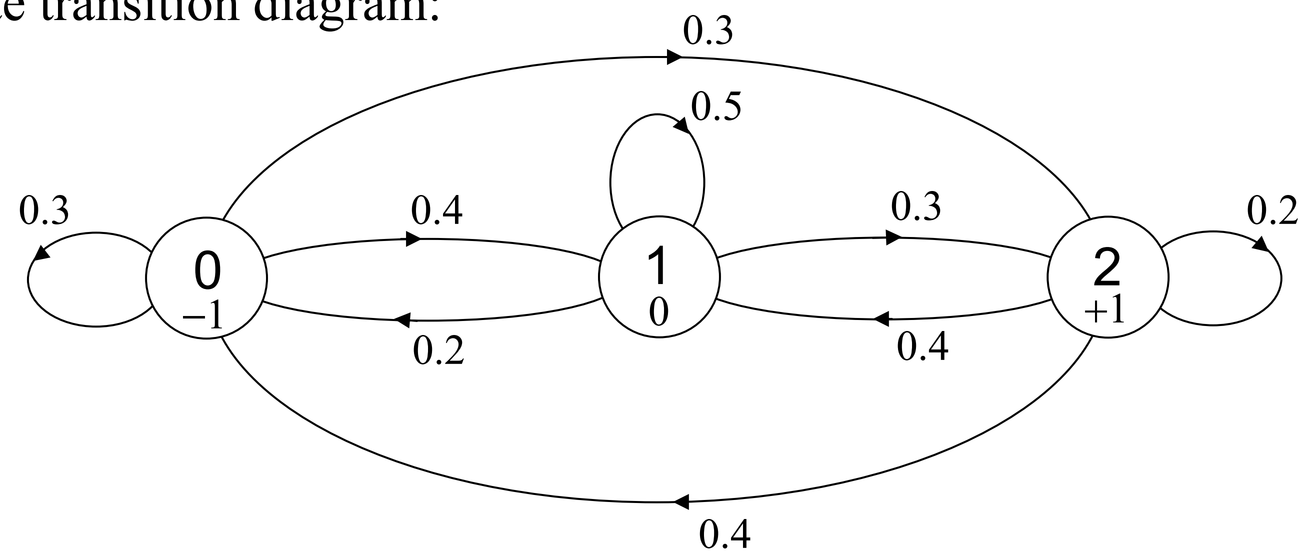
State transition probabilities:

$$p_{00} = 0.3 \qquad p_{01} = 0.4$$
$$p_{10} = 0.2 \qquad p_{11} = 0.5$$
$$p_{20} = 0.4 \qquad p_{21} = 0.4$$

State transition matrix:

$$\mathbf{P} = \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.4 & 0.2 \end{bmatrix}$$

State transition diagram:



State Probabilities (not transition probabilities)

The state probability vector at any discrete time '*k*' is given by

$$\mathbf{p}[n] = \begin{bmatrix} p_0[n] \\ p_1[n] \\ \vdots \\ p_{m-1}[n] \end{bmatrix} \qquad \sum_{i=0}^{m-1} p_i[n] = 1 \qquad \text{for all } n$$

Then the state vector at time '*k*' is given by

$$\mathbf{p}[n] = \mathbf{P}^T \mathbf{p}[n-1] = \left(\mathbf{P}^T\right)^k \mathbf{p}[0], \qquad n = 1, 2, 3, \cdots$$

Limiting-State Probabilities

The limiting-state probability vector

$$\bar{\mathbf{p}} = \lim_{n\to\infty} \mathbf{p}[n] = \lim_{n\to\infty} \left(\mathbf{P}^T\right)^n \mathbf{p}[0] \quad \Leftarrow \quad \sum_{i=0}^{m-1} \bar{p}_i = 1$$

Assuming that the limiting-state probabilities exist, we have

$$\mathbf{P}^T \bar{\mathbf{p}} = \bar{\mathbf{p}}$$

eigenequation
- eigenvalue = 1
- eigenvector = $\bar{\mathbf{p}}$

$$\left(\mathbf{P}^T - \mathbf{I}\right)\bar{\mathbf{p}} = \mathbf{0}$$

Solve for $\bar{\mathbf{p}}$ (**P** is known).

Example:

$$\mathbf{P} = \begin{bmatrix} 0.6 & 0.4 \\ 0.1 & 0.9 \end{bmatrix}, \qquad \sum_{i=0}^{1} \overline{p}_i = 1$$

(a) Find $\overline{\mathbf{p}}$.

$$\left(\mathbf{P}^T - \mathbf{I}\right)\overline{\mathbf{p}} = \underline{0} \Rightarrow \begin{bmatrix} -0.4 & 0.1 \\ 0.4 & -0.1 \end{bmatrix} \begin{bmatrix} \overline{p}_0 \\ \overline{p}_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

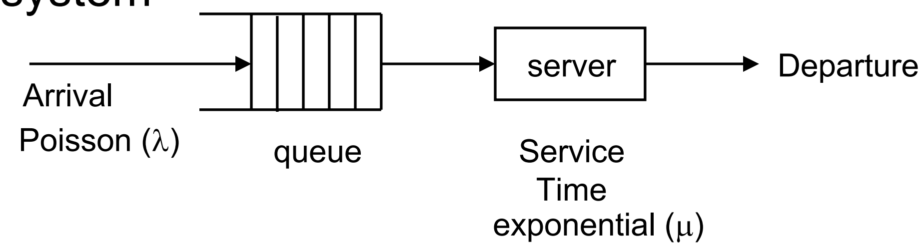
$$\left.\begin{aligned} 4\overline{p}_0 &= \overline{p}_1 \\ \overline{p}_0 + \overline{p}_1 &= 1 \end{aligned}\right\} \Rightarrow \begin{aligned} \overline{p}_0 &= 0.2 \\ \overline{p}_1 &= 0.8 \end{aligned}$$

(b) Find the probability of a run of ten values of state 0

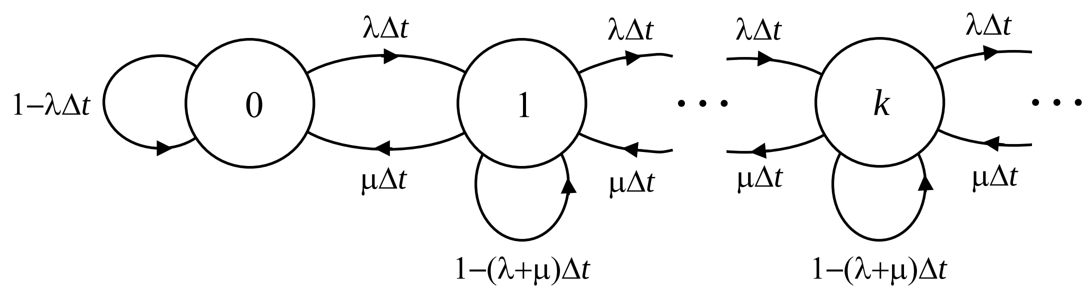
$$(0.2)(0.6)^9 = 0.002$$

Continuous-Time Markov Chain

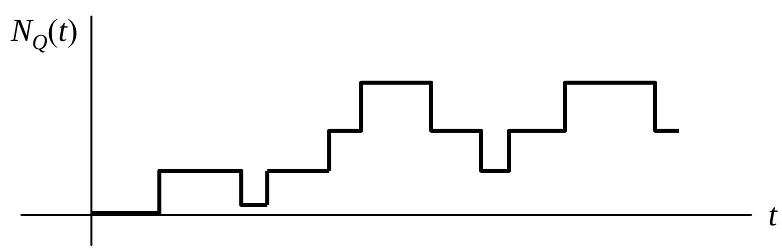
Simple server system



The queue can be described by a continuous-time Markov chain with a possibly *infinite* number of states.

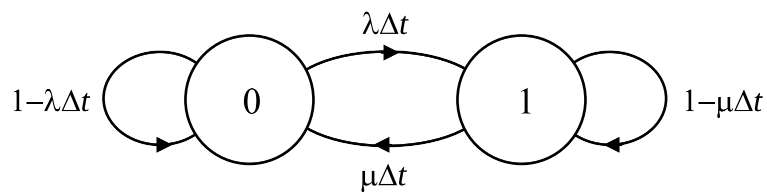


The Markov random process looks like:

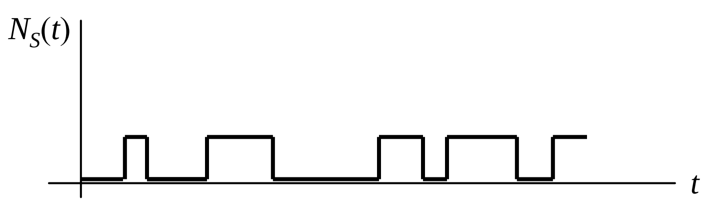


Continuous-Time Markov Chain (cont'd.)

The server can be described by a continuous-time Markov chain with just two states.



The Markov random process looks like:



Transition Rate Diagram

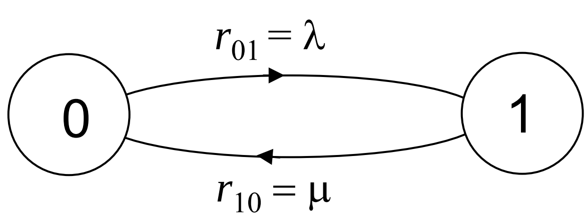
Let $N(t)$ be a continuous-time Markov chain.

<u>*Transition Rates*</u>

r_{ij} are rates of the Poisson process defining transitions from state i to state j.

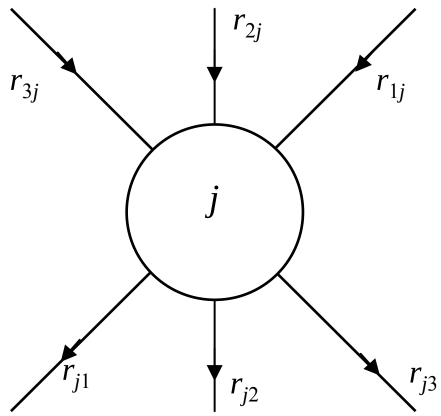
A <u>transition rate diagram</u> is a state diagram with the rates indicated. Note that the rates do not represent probabilities directly and there are no self-loops.

Example for the service process:



Time-Dependent State Probabilities

Now consider a more general Markov chain. A typical state has multiple transitions to and from other states:



Define $p_j(t) = \Pr[N(t) = j]$. Then,

$$p_j(t+\Delta t) = \sum_{\substack{\ell=0 \\ \ell \neq j}}^{m-1} r_{\ell j}\, \Delta t \cdot p_\ell(t) + \left(1 - \sum_{\substack{\ell=0 \\ \ell \neq j}}^{m-1} r_{j\ell}\, \Delta t\right) p_j(t)$$

Time-Dependent State Probabilities (cont'd.)

$$p_j(t+\Delta t) = \sum_{\substack{\ell=0 \\ \ell \neq j}}^{m-1} r_{\ell j}\,\Delta t \cdot p_\ell(t) + \left(1 - \sum_{\substack{\ell=0 \\ \ell \neq j}}^{m-1} r_{j\ell}\,\Delta t\right) p_j(t)$$

Rearranging:

$$\frac{p_j(t+\Delta t) - p_j(t)}{\Delta t} = \sum_{\substack{\ell=0 \\ \ell \neq j}}^{m-1} r_{\ell j}\, p_\ell(t) - \left(\sum_{\substack{\ell=0 \\ \ell \neq j}}^{m-1} r_{j\ell}\right) p_j(t)$$

In the limit:

$$\boxed{\frac{dp_j(t)}{dt} = \sum_{\ell=0}^{m-1} r_{\ell j}\, p_\ell(t)} \quad \text{where} \quad r_{jj} = -\sum_{\substack{\ell=0 \\ \ell \neq j}}^{m-1} r_{j\ell}$$

(Chapman-Kolmogorov equation for continuous-time Markov chains)

Global Balance Equations

Beginning with the Chapman-Kolmogorov equation:

$$\frac{dp_j(t)}{dt} = \sum_{\ell=0}^{m-1} r_{\ell j}\, p_\ell(t) \qquad j = 0,1,\cdots,m-1$$

For steady-state probability flow-rate balance, let:

$$t \to \infty \;\Rightarrow\; p_j(t) \to p_j \;\Rightarrow\; \frac{dp_j(t)}{dt} \to 0$$

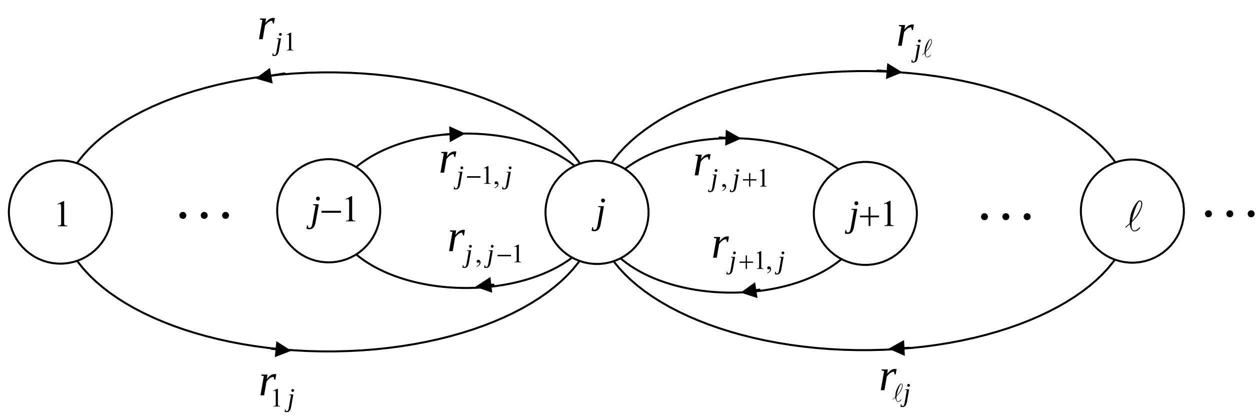
$$\therefore \quad \sum_{\ell=0}^{m-1} r_{\ell j}\, p_\ell = 0 \qquad j = 0,1,\cdots,m-1$$

Global Balance Equations (cont'd.)

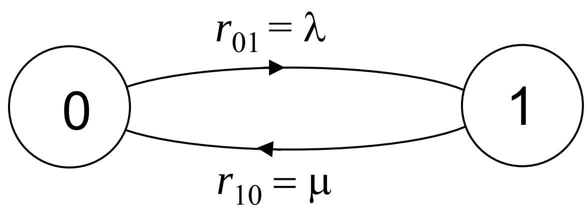
$$\sum_{\ell=0}^{m-1} r_{\ell j}\, p_\ell = 0 \qquad j = 0,1,\cdots,m-1$$

Now recall that, $r_{jj} = -\sum_{\substack{\ell=0 \\ \ell\neq j}}^{m-1} r_{j\ell}$ Thus $\sum_{\substack{\ell=0 \\ \ell\neq j}}^{m-1} r_{\ell j}\, p_\ell - p_j \sum_{\substack{\ell=0 \\ \ell\neq j}}^{m-1} r_{j\ell} = 0$

Probability flow-rate balance equation: $\boxed{p_j \sum_{\substack{\ell=0 \\ \ell\neq j}}^{m-1} r_{j\ell} = \sum_{\substack{\ell=0 \\ \ell\neq j}}^{m-1} r_{\ell j}\, p_\ell}$



<u>Example</u>:



For State 0: $\lambda p_0 = \mu p_1$

(Global balance equations)

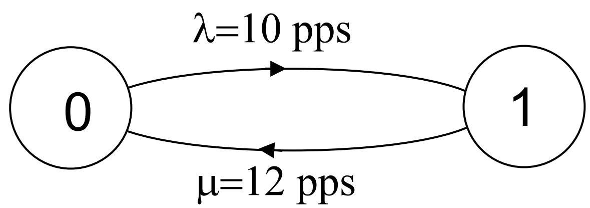
For State 1: $\mu p_1 = \lambda p_0$

Also use: $p_0 + p_1 = 1$

to find:

$$p_0 = \frac{\mu}{\lambda + \mu}; \qquad p_1 = \frac{\lambda}{\lambda + \mu}$$

Numerical Example:

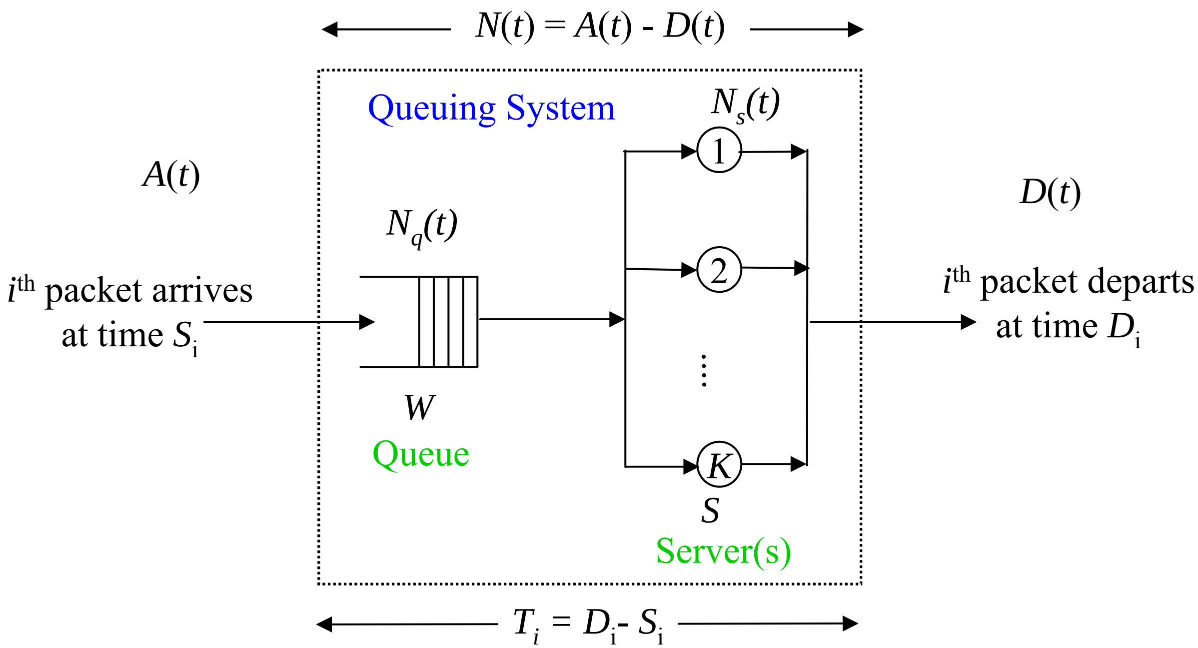


$$\lambda p_0 = \mu p_1 \quad \Rightarrow \quad p_1 = \frac{\lambda}{\mu} p_0 = \frac{5}{6} p_0$$

$$p_0 + p_1 = 1 \quad \Rightarrow \quad p_0 + \frac{5}{6} p_0 = 1$$

$$p_0 = \frac{6}{11}, \quad p_1 = \frac{5}{11}$$

Queuing System: M/M/K†



$N(t)$: Number of packets in system

$N_q(t)$: Number of packets in queue

$N_s(t)$: Number of packets in service

T: Time spent in system

W: Waiting time in queue

S: Service time

† Shorthand notation: $a/b/c/d$

Arrival process/Service time distribution/Number of servers/Buffer size

Single Server Queuing System: M/M/1 System

- Job arrival is a Poisson process and the interarrival time τ is exponentially distributed with λ as the parameter

$$\Pr\left[A(t)=n\right]=\frac{(\lambda t)^n}{n!}e^{-\lambda t}, \qquad n=0,1,2,\cdots$$

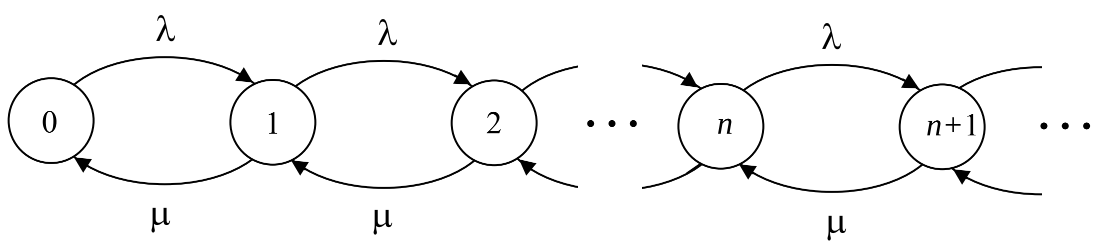
$$f_\tau(\tau)=\lambda e^{-\lambda\tau}, \qquad \tau\geq 0$$

- Job service time S is exponentially distributed with μ as the parameter

$$f_S(s)=\mu e^{-\mu s}, \qquad s\geq 0$$

M/M/1 System (cont'd.)

Assume an infinitely long queue (birth-death queue)



with state transition rates: $r_{j,j+1} = \lambda, \qquad r_{j,j-1} = \mu, \qquad j = 0,1,2,\cdots$

<u>Global Balance Equations</u> (by observation)

State 0:	$\lambda p_0 = \mu p_1$	$\lambda p_0 - \mu p_1 = 0$
State 1:	$(\lambda + \mu) p_1 = \lambda p_0 + \mu p_2$	$\lambda p_1 - \mu p_2 = \lambda p_0 - \mu p_1 = 0$
$\cdots$		$\cdots$
State n:	$(\lambda + \mu) p_n = \lambda p_{n-1} + \mu p_{n+1}$	$\lambda p_n - \mu p_{n+1} = \lambda p_{n-1} - \mu p_n = 0$

M/M/1 System (cont'd.)

<u>Global balance equations</u>

$$\lambda p_0 - \mu p_1 = 0$$

$$\cdots$$

$$\lambda p_1 - \mu p_2 = 0$$

$$\lambda p_n - \mu p_{n+1} = 0 \qquad \text{or} \quad p_j = \frac{\lambda}{\mu} p_{j-1} \qquad j = 1, 2, 3, \cdots, n, n+1, \cdots$$

By induction, the state probabilities:

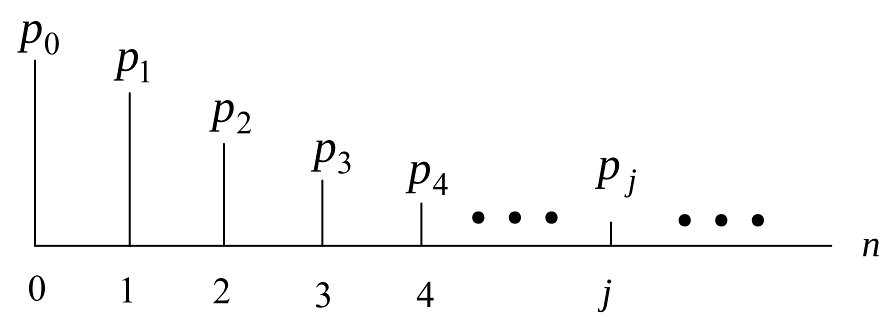
$$\boxed{p_n = \left(\frac{\lambda}{\mu}\right)^n p_0; \qquad n = 0, 1, 2, 3, \cdots}$$

M/M/1 System (cont'd.)

Rewrite: $$\boxed{p_n = \rho^n p_0; \qquad n = 0, 1, 2, 3, \cdots \quad \text{where} \;\; \rho = \frac{\lambda}{\mu}}$$

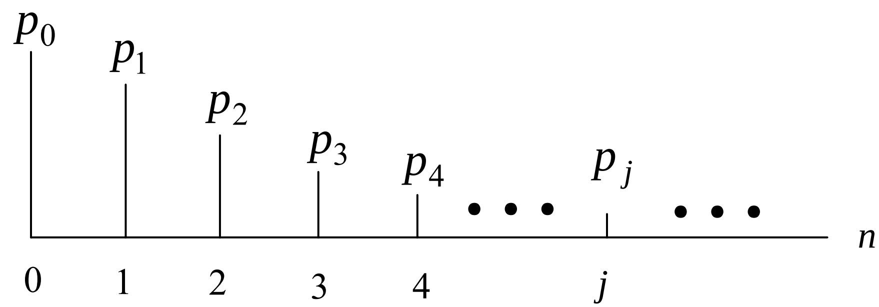
This defines a PMF for the state of the queue

$$\mathrm{f}_N[n] = \Pr[N = n] = p_n$$



M/M/1 System (cont'd.)

$$\mathrm{f}_N\left[n\right] = \Pr\left[N = n\right] = p_n$$



To determine the unknown p_0 use:

$$\sum_{j=0}^{\infty} p_j = \sum_{j=0}^{\infty} \rho^j p_0 \qquad p_0 \sum_{j=0}^{\infty} \rho^j = p_0 \frac{1}{1-\rho} = 1$$

for $\rho < 1$, i.e., $\lambda < \mu$, thus $p_0 = 1 - \rho$ and

$$\boxed{\mathrm{f}_N\left[n\right] = p_n = \left(1-\rho\right)\rho^n, \qquad n = 0, 1, 2, \cdots}$$

Thus N, which represents the number of jobs in the system, is a type 0 *geometric* random variable. ρ is called the ***utilization factor*** $(0 < \rho < 1)$.

Some Important Formulas

1. Probability of long queues:

$$\Pr\left[N < n_0\right] = \sum_{j=0}^{n_0-1} (1-\rho)\rho^j = (1-\rho)\sum_{j=0}^{n_0-1} \rho^j$$

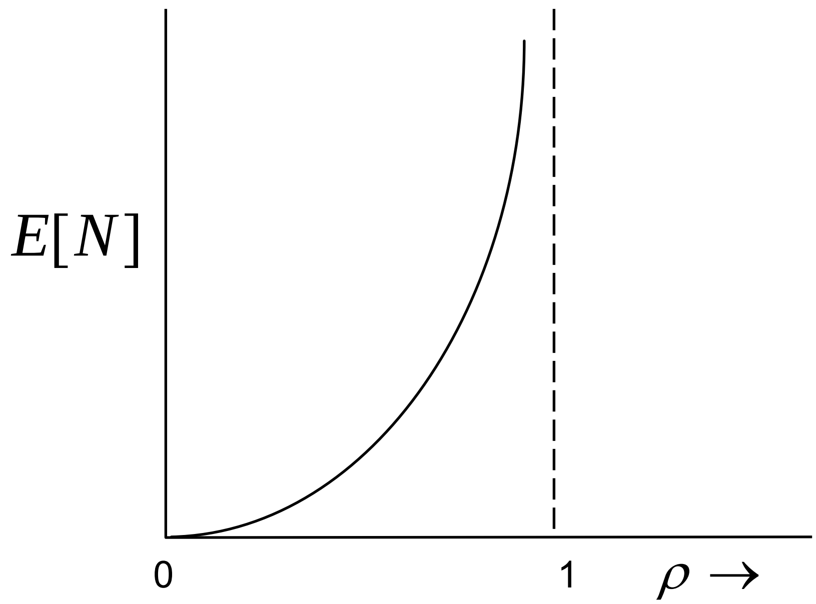
$$= (1-\rho)\cdot\frac{1-\rho^{n_0}}{1-\rho} = 1-\rho^{n_0}$$

$$\boxed{\Pr\left[N \geq n_0\right] = \rho^{n_0}}$$

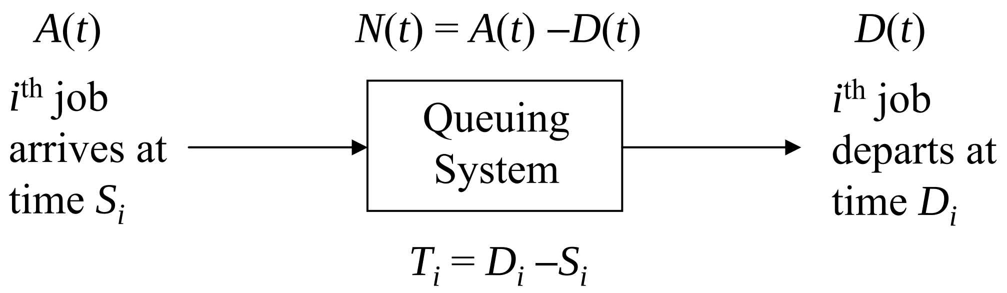
Important Formulas (cont'd.)

2. Average number of jobs in the system:

$$\boxed{E\left[N(t)\right] = \frac{\rho}{1-\rho}}$$ (mean of the geometric PMF)



Little's Formula



For systems that reach equilibrium, the average number of jobs N in a system is

$$E[N(t)] = \lambda\, E[T]$$

where $E[T]$ is the average time spent in the system by a job.

Comments on Little's Formula

Little's formula can be extended to other quantities. Let $T = W + S$

where *W* is the waiting time in the queue and *S* is the service time (note that $E[S] = 1/\mu$).

- The average number of jobs *in the queue* is given by

$$E\left[N_q(t)\right] = \lambda E[W] \quad (1)$$

- The average number of jobs in service or utilization of a single server system is

$$E\left[N_S(t)\right] = \lambda E[S] = \lambda / \mu = \rho \quad (2)$$

  where $E[S]$ is the average service time.

- The utilization of a *K*-server system is then given by

$$\rho_K = \frac{\lambda E[S]}{K} = \frac{\lambda}{K\mu}$$

Comments (cont'd.)

- The average total delay experienced by a job in a single server system

$$E[T] = \frac{E[N(t)]}{\lambda} \quad \leftarrow \text{from little's formula}$$

$$= \frac{1}{\lambda} \cdot \frac{\rho}{1-\rho} = \frac{1}{\lambda} \cdot \frac{\lambda/\mu}{1-\lambda/\mu} \quad \leftarrow \text{Note} \quad \rho = \frac{\lambda}{\mu}$$

$$E[T] = \frac{1}{\mu - \lambda} \tag{3}$$

Comments (cont'd.)

- The average waiting time in the queue

$$E[W] = E[T] - E[S] = \frac{\rho}{\lambda(1-\rho)} - \frac{\rho}{\lambda} = \frac{\rho^2}{\lambda(1-\rho)} \qquad (4)$$

- The average number of jobs in the queue

$$E[N_q] = \lambda E[W] \qquad \leftarrow \text{Little's formula restated}$$

$$= \frac{\rho^2}{1-\rho} \qquad (5)$$

Example:

Consider an M/M/1 system

(a) Find Pr [$N(t) > 10$]

$$\Pr\left[N(t) > 10\right] = \Pr\left[N(t) \geq 11\right] = \rho^{11}$$

(b) Find the maximum allowable arrival rate, λ, if we require $\Pr\left[N(t) \geq 10\right] = 10^{-3}$. Let $\mu = 4$ per second

$$\Pr\left[N(t) \geq 10\right] = \rho^{10} = 10^{-3}$$

$$\therefore \quad \rho = \frac{\lambda}{\mu} = 10^{-0.3}$$

$$\lambda = \mu 10^{-0.3} \cong 2$$

Example (cont'd.):

(c) Find the minimum allowable service rate, μ, if we require $\Pr\left[N(t)\geq 10\right]=10^{-5}$. Let $\lambda = 4$ per second

$$\Pr\left[N(t)\geq 10\right]=\rho^{10}=10^{-5}$$

$$\therefore \quad \rho=\frac{\lambda}{\mu}=10^{-0.5}$$

$$\mu=\frac{\lambda}{10^{-0.5}}\cong 13$$

Example: M/M/1 system

Packets arrive at a network router at a rate $\lambda = 2\times10^5$ per sec.

The service is performed by the router at a rate $\mu = 2.5\times10^5$ per sec.

The utilization factor, $\rho = \dfrac{\lambda}{\mu} = \dfrac{2}{2.5} = \dfrac{4}{5}$

PMF, $\quad \mathrm{f}_N[n] = (1-\rho)\rho^n,\ n \geq 0.$

(a) Find the mean number of jobs in the system, $E[N(t)]$.

$$E\left[N(t)\right] = \frac{\rho}{1-\rho} = \frac{0.8}{1-0.8} = 4$$

Example (cont'd.):

(b) Average total delay in the system. From Little's formula

$$E[T] = \frac{E[N(t)]}{\lambda} = \frac{4}{2\times 10^5} = 20\ \mu\text{s}$$

(c) What value of λ would double $E[T]$ in (b)?

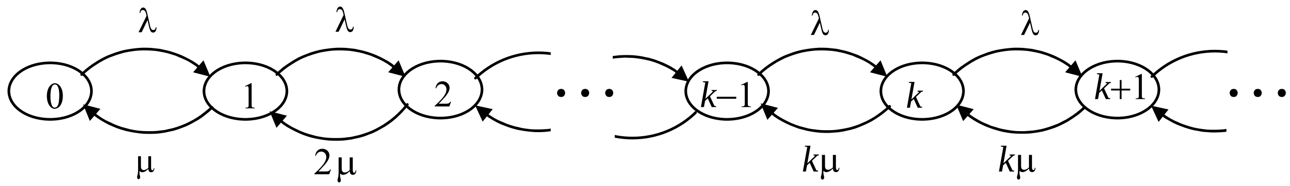
$$E[T] = \frac{1}{\mu-\lambda} = 40\ \mu\text{s} \quad \text{or} \quad \mu-\lambda = \frac{10^6}{40} = 0.25\times 10^5$$

$$\lambda = \mu - 0.25\times 10^5 = 2.5\times 10^5 - 0.25\times 10^5 = 2.25\times 10^5 \ \text{ per sec}$$

(d) What is the utilization? (0.9)

M/M/k System

The state transition diagram for an M/M/k system



The state probabilities of this queue can be written as

$$p_n = \frac{\lambda}{n\mu} p_{n-1} = \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n p_0, \qquad n = 1, 2, \cdots, k$$

$$p_n = \frac{\lambda}{k\mu} p_{n-1} = \left(\frac{\lambda}{k\mu}\right)^{n-k} p_n = \rho^{n-k} p_k, \qquad n = k+1, k+2, \cdots$$

where ρ is defined as $\rho = \lambda/k\mu$.

M/M/k System (cont'd.)

By substituting for p_k, we have

$$p_n = \frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k \rho^{n-k} p_0, \qquad n = k+1,\ k+2,\ k+3,\cdots$$

The probability of the 0th state is obtained as follows:

$$\sum_{n=0}^{\infty} p_n = 1 = p_0 + p_0 \sum_{n=1}^{k-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n + p_0 \sum_{n=k}^{\infty} \frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k \rho^{n-k}$$

$$p_0 = \left[\sum_{n=0}^{k-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n + \frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k \frac{1}{1-\rho}\right]^{-1}$$

M/M/k System (cont'd.)

The probability that an arriving job is forced to wait because all *k*-servers are busy is known as the Erlang C formula:

$$C\left(k, \frac{\lambda}{\mu}\right) = \Pr\left[W > 0\right] = \Pr\left[N > k\right] = \sum_{n=k}^{\infty} p_n$$

$$= \sum_{n=k}^{\infty} \rho^{n-k}\, p_k = \frac{p_k}{1-\rho}$$

The average number of customers in the queue is given by

$$E\left[N_q\right] = \sum_{n=k}^{\infty} (n-k)\,\rho^{n-k} p_k = p_k \sum_{i=0}^{\infty} i\rho^i = \frac{\rho}{(1-\rho)^2}\, p_k$$

M/M/k System (cont'd.)

From Little's formula: $$E[W] = \frac{E[N_q]}{\lambda} = \frac{\rho}{\lambda(1-\rho)^2} p_k$$

The system delay is then given by:

$$E[T] = E[W] + E[S] = \frac{\rho}{\lambda(1-\rho)^2} p_k + \frac{1}{\mu}$$

$$= \frac{1}{k!} \frac{\rho}{\lambda(1-\rho)^2} \left(\frac{\lambda}{\mu}\right)^k p_0 + \frac{1}{\mu}$$

By using Little's formula, the total jobs in the system can then be obtained as

$$E[N] = \lambda E[T]$$

- Continuous-time Markov chain
  - M/M/1 queue

$$\boxed{E\left[N(t)\right]=\frac{\rho}{1-\rho}} \qquad E\left[N(t)\right]=\lambda E\left[T\right] \qquad \rho = \lambda/\mu$$

  - M/M/k queue

$$E\left[T\right]=\frac{1}{k!}\frac{\rho}{\lambda(1-\rho)^2}\left(\frac{\lambda}{\mu}\right)^k p_0+\frac{1}{\mu} \qquad \rho = \lambda/k\mu$$

Example:

Compare queuing systems:

A. M/M/1 system with one 100 GIPS server

$$\lambda = 5000 \text{ per sec}, \ \ \mu = 6000 \text{ per sec}$$

$$E[T] = \frac{1}{\mu - \lambda} = \frac{1}{6000 - 5000} = 1\,\text{ms}$$

B. Ten M/M/1 systems each with a 10 GIPS server

with $\lambda = 500$ per sec, $\mu = 600$ per sec

$$E[T] = \frac{1}{\mu - \lambda} = \frac{1}{600 - 500} = 10\,\text{ms}$$

The mean delay in the 100 GIPS system is 1/10$^{\text{th}}$ that of the system with ten 10 GIPS servers.

Example (cont'd.):

C. An M/M/10 system with a 10 GIPS server:

$$E\left[T\right] = \frac{1}{k!} \frac{\rho}{\lambda(1-\rho)^2} \left(\frac{\lambda}{\mu}\right)^k p_0 + \frac{1}{\mu}$$

$$= 8.8278 \times 10^{-12} p_0 + \frac{1}{\mu} \cong \frac{1}{\mu}$$

$$= \frac{1}{600} = 1.6667 \text{ ms}$$